

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/566,606
Source: PCR
Date Processed by STIC: 2/7/06

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/10/566,606

TIME: 09:51:17

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\02012006\J566606.raw

```

5 <110> APPLICANT: Degussa AG
8 <120> TITLE OF INVENTION: Process for the preparation of L-threonine
11 <130> FILE REFERENCE: 030217BT
C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/566,606
C--> 14 <141> CURRENT FILING DATE: 2006-01-31
14 <160> NUMBER OF SEQ ID NOS: 10
17 <170> SOFTWARE: PatentIn version 3.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 993
22 <212> TYPE: DNA
23 <213> ORGANISM: Escherichia coli
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (1)..(990)
29 <223> OTHER INFORMATION: rpoS gene
32 <400> SEQUENCE: 1
33 atg agt cag aat acg ctg aaa gtt cat gat tta aat gaa gat gcg gaa      48
34 Met Ser Gln Asn Thr Leu Lys Val His Asp Leu Asn Glu Asp Ala Glu
35 1          5          10          15
37 ttt gat gag aac gga gtt gag gtt ttt gac gaa aag gcc tta gta gaa      96
38 Phe Asp Glu Asn Gly Val Glu Val Phe Asp Glu Lys Ala Leu Val Glu
39          20          25          30
41 cag gaa ccc agt gat aac gat ttg gcc gaa gag gaa ctg tta tcg cag      144
42 Gln Glu Pro Ser Asp Asn Asp Leu Ala Glu Glu Glu Leu Leu Ser Gln
43          35          40          45
45 gga gcc aca cag cgt gtg ttg gac gcg act cag ctt tac ctt ggt gag      192
46 Gly Ala Thr Gln Arg Val Leu Asp Ala Thr Gln Leu Tyr Leu Gly Glu
47          50          55          60
49 att ggt tat tca cca ctg tta acg gcc gaa gaa gaa gtt tat ttt gcg      240
50 Ile Gly Tyr Ser Pro Leu Leu Thr Ala Glu Glu Glu Val Tyr Phe Ala
51 65          70          75          80
53 cgt cgc gca ctg cgt gga gat gtc gcc tct cgc cgc cgg atg atc gag      288
54 Arg Arg Ala Leu Arg Gly Asp Val Ala Ser Arg Arg Arg Met Ile Glu
55          85          90          95
57 agt aac ttg cgt ctg gtg gta aaa att gcc cgc cgt tat ggc aat cgt      336
58 Ser Asn Leu Arg Leu Val Val Lys Ile Ala Arg Arg Tyr Gly Asn Arg
59          100         105         110
62 ggt ctg gcg ttg ctg gac ctt atc gaa gag ggc aac ctg ggg ctg atc      384
63 Gly Leu Ala Leu Leu Asp Leu Ile Glu Glu Gly Asn Leu Gly Leu Ile
64          115         120         125
66 cgc gcg gta gag aag ttt gac ccg gaa cgt ggt ttc cgc ttc tca aca      432
67 Arg Ala Val Glu Lys Phe Asp Pro Glu Arg Gly Phe Arg Phe Ser Thr
68          130         135         140

```

RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/10/566,606

TIME: 09:51:17

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\02012006\J566606.raw

```

70 tac gca acc tgg tgg att cgc cag acg att gaa cgg gcg att atg aac      480
71 Tyr Ala Thr Trp Trp Ile Arg Gln Thr Ile Glu Arg Ala Ile Met Asn
72 145                      150                      155                      160
74 caa acc cgt act att cgt ttg ccg att cac atc gta aag gag ctg aac      528
75 Gln Thr Arg Thr Ile Arg Leu Pro Ile His Ile Val Lys Glu Leu Asn
76                      165                      170                      175
78 gtt tac ctg cga acc gca cgt gag ttg tcc cat aag ctg gac cat gaa      576
79 Val Tyr Leu Arg Thr Ala Arg Glu Leu Ser His Lys Leu Asp His Glu
80                      180                      185                      190
82 cca agt gcg gaa gag atc gca gag caa ctg gat aag cca gtt gat gac      624
83 Pro Ser Ala Glu Glu Ile Ala Glu Gln Leu Asp Lys Pro Val Asp Asp
84                      195                      200                      205
86 gtc agc cgt atg ctt cgt ctt aac gag cgc att acc tcg gta gac acc      672
87 Val Ser Arg Met Leu Arg Leu Asn Glu Arg Ile Thr Ser Val Asp Thr
88                      210                      215                      220
90 ccg ctg ggt ggt gat tcc gaa aaa gcg ttg ctg gac atc ctg gcc gat      720
91 Pro Leu Gly Gly Asp Ser Glu Lys Ala Leu Leu Asp Ile Leu Ala Asp
92 225                      230                      235                      240
94 gaa aaa gag aac ggt ccg gaa gat acc acg caa gat gac gat atg aag      768
95 Glu Lys Glu Asn Gly Pro Glu Asp Thr Thr Gln Asp Asp Asp Met Lys
96                      245                      250                      255
98 cag agc atc gtc aaa tgg ctg ttc gag ctg aac gcc aaa cag cgt gaa      816
99 Gln Ser Ile Val Lys Trp Leu Phe Glu Leu Asn Ala Lys Gln Arg Glu
100                      260                      265                      270
102 gtg ctg gca cgt cga ttc ggt ttg ctg ggg tac gaa gcg gca aca ctg      864
103 Val Leu Ala Arg Arg Phe Gly Leu Leu Gly Tyr Glu Ala Ala Thr Leu
104                      275                      280                      285
106 gaa gat gta ggt cgt gaa att ggc ctc acc cgt gaa cgt gtt cgc cag      912
107 Glu Asp Val Gly Arg Glu Ile Gly Leu Thr Arg Glu Arg Val Arg Gln
108                      290                      295                      300
110 att cag gtt gaa ggc ctg cgc cgt ttg cgc gaa atc ctg caa acg cag      960
111 Ile Gln Val Glu Gly Leu Arg Arg Leu Arg Glu Ile Leu Gln Thr Gln
112 305                      310                      315                      320
114 ggg ctg aat atc gaa gcg ctg ttc cgc gag taa      993
115 Gly Leu Asn Ile Glu Ala Leu Phe Arg Glu
116                      325                      330
118 <210> SEQ ID NO: 2
119 <211> LENGTH: 330
120 <212> TYPE: PRT
121 <213> ORGANISM: Escherichia coli
124 <400> SEQUENCE: 2
125 Met Ser Gln Asn Thr Leu Lys Val His Asp Leu Asn Glu Asp Ala Glu
126 1                      5                      10                      15
128 Phe Asp Glu Asn Gly Val Glu Val Phe Asp Glu Lys Ala Leu Val Glu
129                      20                      25                      30
131 Gln Glu Pro Ser Asp Asn Asp Leu Ala Glu Glu Glu Leu Leu Ser Gln
132                      35                      40                      45
134 Gly Ala Thr Gln Arg Val Leu Asp Ala Thr Gln Leu Tyr Leu Gly Glu
135                      50                      55                      60

```

RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/10/566,606

TIME: 09:51:17

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\02012006\J566606.raw

```

137 Ile Gly Tyr Ser Pro Leu Leu Thr Ala Glu Glu Glu Val Tyr Phe Ala
138 65          70          75          80
140 Arg Arg Ala Leu Arg Gly Asp Val Ala Ser Arg Arg Arg Met Ile Glu
141          85          90          95
143 Ser Asn Leu Arg Leu Val Val Lys Ile Ala Arg Arg Tyr Gly Asn Arg
144          100         105         110
146 Gly Leu Ala Leu Leu Asp Leu Ile Glu Glu Gly Asn Leu Gly Leu Ile
147          115         120         125
149 Arg Ala Val Glu Lys Phe Asp Pro Glu Arg Gly Phe Arg Phe Ser Thr
150          130         135         140
152 Tyr Ala Thr Trp Trp Ile Arg Gln Thr Ile Glu Arg Ala Ile Met Asn
153 145          150         155         160
155 Gln Thr Arg Thr Ile Arg Leu Pro Ile His Ile Val Lys Glu Leu Asn
156          165         170         175
158 Val Tyr Leu Arg Thr Ala Arg Glu Leu Ser His Lys Leu Asp His Glu
159          180         185         190
161 Pro Ser Ala Glu Glu Ile Ala Glu Gln Leu Asp Lys Pro Val Asp Asp
162          195         200         205
164 Val Ser Arg Met Leu Arg Leu Asn Glu Arg Ile Thr Ser Val Asp Thr
165          210         215         220
167 Pro Leu Gly Gly Asp Ser Glu Lys Ala Leu Leu Asp Ile Leu Ala Asp
168 225          230         235         240
170 Glu Lys Glu Asn Gly Pro Glu Asp Thr Thr Gln Asp Asp Asp Met Lys
171          245         250         255
173 Gln Ser Ile Val Lys Trp Leu Phe Glu Leu Asn Ala Lys Gln Arg Glu
174          260         265         270
176 Val Leu Ala Arg Arg Phe Gly Leu Leu Gly Tyr Glu Ala Ala Thr Leu
177          275         280         285
179 Glu Asp Val Gly Arg Glu Ile Gly Leu Thr Arg Glu Arg Val Arg Gln
180          290         295         300
182 Ile Gln Val Glu Gly Leu Arg Arg Leu Arg Glu Ile Leu Gln Thr Gln
183 305          310         315         320
186 Gly Leu Asn Ile Glu Ala Leu Phe Arg Glu
187          325         330
189 <210> SEQ ID NO: 3
190 <211> LENGTH: 993
191 <212> TYPE: DNA
192 <213> ORGANISM: Escherichia coli
195 <220> FEATURE:
196 <221> NAME/KEY: Allele
197 <222> LOCATION: (1)..(990)
198 <223> OTHER INFORMATION: rpoS allele
201 <220> FEATURE:
202 <221> NAME/KEY: misc_feature
203 <222> LOCATION: (97)..(99)
204 <223> OTHER INFORMATION: amber codon
207 <400> SEQUENCE: 3
208 atgagtcaga atacgctgaa agttcatgat ttaaataaag atgcggaatt tgatgagaac      60
210 ggagttgagg tttttgacga aaaggcetta gtagaatagg aaccagtgga taacgatttg      120

```

RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/10/566,606

TIME: 09:51:17

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\02012006\J566606.raw

```

212 gccgaagagg aactgttata gcagggagcc acacagcgtg tgttggacgc gactcagctt 180
214 taccttggtg agattgggta ttcaccactg ttaacggccg aagaagaagt ttattttgcg 240
216 cgtcgcgcac tgcgtggaga tgcgcctct cgccgccgga tgatcgagag taacttgcgt 300
218 ctggtggtaa aaattgccc cgttatggc aatcgtggtc tggcgttgct ggacctatc 360
220 gaagagggca acctggggct gatccgcgcg gtagagaagt ttgaccgga acgtggtttc 420
222 cgcttctcaa catacgcaac ctggtggatt cgccagacga ttgaacgggc gattatgaac 480
224 caaaccgta ctattcggtt gccgattcac atcgtaaagg agctgaacgt ttacctgcga 540
226 accgcacgtg agttgtccca taagctggac catgaaccaa gtgcggaaga gatcgagag 600
228 caactggata agccagttga tgacgtcagc cgtatgcttc gtcttaacga gcgcattacc 660
230 tcggtagaca ccccgctggg tgggtgattcc gaaaaagcgt tgctggacat cctggccgat 720
232 gaaaaagaga acggtccgga agataccacg caagatgacg atatgaagca gagcatcgtc 780
234 aaatggctgt tcgagctgaa cgccaaacag cgtgaagtgc tggcacgtcg attcggtttg 840
236 ctgggggtac aagcggcaac actggaagat gtaggtcgtg aaattggcct caccggtgaa 900
238 cgtgttcgcc agattcaggt tgaaggcctg cgccgtttgc gcgaaatcct gcaaacgcag 960
240 gggtgaata tcgaagcgt gttccgcgag taa 993
242 <210> SEQ ID NO: 4
243 <211> LENGTH: 75
244 <212> TYPE: DNA
245 <213> ORGANISM: Escherichia coli
248 <220> FEATURE:
249 <221> NAME/KEY: tRNA
250 <222> LOCATION: (1)..(75)
251 <223> OTHER INFORMATION: supE allele
254 <400> SEQUENCE: 4
255 tggggatatcg ccaagcggta aggcaccgga ttctaattcc ggcattccga ggttcgaatc 60
257 ctctaccccc agcca 75
259 <210> SEQ ID NO: 5
260 <211> LENGTH: 1545
261 <212> TYPE: DNA
262 <213> ORGANISM: Escherichia coli
265 <220> FEATURE:
266 <221> NAME/KEY: CDS
267 <222> LOCATION: (1)..(1542)
268 <223> OTHER INFORMATION: ilvA gene
270 <400> SEQUENCE: 5
271 atg gct gac tcg caa ccc ctg tcc ggt gct ccg gaa ggt gcc gaa tat 48
272 Met Ala Asp Ser Gln Pro Leu Ser Gly Ala Pro Glu Gly Ala Glu Tyr
273 1 5 10 15
275 tta aga gca gtg ctg cgc gcg ccg gtt tac gag gcg gcg cag gtt acg 96
276 Leu Arg Ala Val Leu Arg Ala Pro Val Tyr Glu Ala Ala Gln Val Thr
277 20 25 30
279 ccg cta caa aaa atg gaa aaa ctg tcg tcg cgt ctt gat aac gtc att 144
280 Pro Leu Gln Lys Met Glu Lys Leu Ser Ser Arg Leu Asp Asn Val Ile
281 35 40 45
283 ctg gtg aag cgc gaa gat cgc cag cca gtg cac agc ttt aag ctg cgc 192
284 Leu Val Lys Arg Glu Asp Arg Gln Pro Val His Ser Phe Lys Leu Arg
285 50 55 60
287 ggc gca tac gcc atg atg gcg ggc ctg acg gaa gaa cag aaa gcg cac 240
288 Gly Ala Tyr Ala Met Met Ala Gly Leu Thr Glu Glu Gln Lys Ala His

```

RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/10/566,606

TIME: 09:51:17

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\02012006\J566606.raw

289	65		70		75		80	
291	ggc	gtg	atc	act	gct	tct	gcg	ggt
292	Gly	Val	Ile	Thr	Ala	Ser	Ala	Gly
293				85			90	
295	tct	tct	gcg	cgg	tta	ggc	gtg	aag
296	Ser	Ser	Ala	Arg	Leu	Gly	Val	Lys
297				100			105	
299	acc	gcc	gac	atc	aaa	gtc	gac	gcg
300	Thr	Ala	Asp	Ile	Lys	Val	Asp	Ala
301				115			120	
303	ctg	ctc	cac	ggc	gcg	aac	ttt	gat
304	Leu	Leu	His	Gly	Ala	Asn	Phe	Asp
305				130			135	
307	ctg	tca	cag	cag	cag	ggg	ttc	acc
308	Leu	Ser	Gln	Gln	Gln	Gly	Phe	Thr
309	145						150	
311	atg	gtg	att	gcc	ggg	caa	ggc	acg
312	Met	Val	Ile	Ala	Gly	Gln	Gly	Thr
313				165			170	
315	gac	gcc	cat	ctc	gac	cgc	gta	ttt
316	Asp	Ala	His	Leu	Asp	Arg	Val	Phe
317				180			185	
319	gct	gct	ggc	gtg	gcg	gtg	ctg	atc
320	Ala	Ala	Gly	Val	Ala	Val	Leu	Ile
321				195			200	
323	gtg	atc	gcc	gta	gaa	gcg	gaa	gac
324	Val	Ile	Ala	Val	Glu	Ala	Glu	Asp
325				210			215	
327	gat	gcg	ggt	cat	ccg	ggt	gat	ctg
328	Asp	Ala	Gly	His	Pro	Val	Asp	Leu
329	225						230	
331	ggc	gta	gcg	gta	aaa	cgc	atc	ggt
332	Gly	Val	Ala	Val	Lys	Arg	Ile	Gly
333				245			250	
335	gag	tat	ctc	gac	gac	atc	atc	acc
336	Glu	Tyr	Leu	Asp	Asp	Ile	Ile	Thr
337				260			265	
339	gcg	atg	aag	gat	tta	ttc	gaa	gat
340	Ala	Met	Lys	Asp	Leu	Phe	Glu	Asp
341				275			280	
343	ggc	gcg	ctg	gcg	ctg	gcg	gga	atg
344	Gly	Ala	Leu	Ala	Leu	Ala	Gly	Met
345				290			295	
347	att	cgc	ggc	gaa	cgg	ctg	gcg	cat
348	Ile	Arg	Gly	Glu	Arg	Leu	Ala	His
349	305						310	
351	ttc	cac	ggc	ctg	cgc	tac	gtc	tca
352	Phe	His	Gly	Leu	Arg	Tyr	Val	Ser
353				325			330	

VERIFICATION SUMMARY

DATE: 02/07/2006

PATENT APPLICATION: US/10/566,606

TIME: 09:51:18

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\02012006\J566606.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:529 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:526
L:777 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:9
L:788 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:9,Line#:779